

Final Report for Period: 07/2011 - 09/2011**Submitted on:** 10/05/2011**Principal Investigator:** Vengazhiyil, Roshan J.**Award ID:** 0448774**Organization:** Georgia Tech Research Corp**Submitted By:**

Vengazhiyil, Roshan - Principal Investigator

Title:

CAREER: Design and Analysis of Experiments for Developing Robust Products and Processes

Project Participants

Senior Personnel

Name: Vengazhiyil, Roshan**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Did research on Bayesian methods for the design and analysis of experiments and supervised two Ph.D. students. Published the project findings in many journals and presented them in many conferences. Two months of support received in Summer 2005, 2006, and 2007 and travel support to attend the conferences.

Post-doc

Graduate Student

Name: Delaney, James**Worked for more than 160 Hours:** Yes**Contribution to Project:**

As part of his Ph.D. thesis, Jim worked on an important topic of this project. One of the major contributions of this project is in developing a prior distribution for a linear model's parameters using Gaussian processes. I have developed this for 2-level experiments. Jim has extended this to multi-level experiments and for qualitative and quantitative factors. He has demonstrated that the proposed Bayesian approach performs better than the existing approaches. His work was partially supported by the grant. He presented the work at the annual meeting of INFORMS 2006 (Pittsburgh), JRC 2006 (Knoxville), DEMA 2006 (Southampton), and NSF DMI Conference (St. Louis) for which the travel expenses were covered by the grant. Jim graduated in Summer 2006 and joined the department of statistics at Carnegie Mellon University as a visiting assistant professor.

Name: Kang, Lulu**Worked for more than 160 Hours:** Yes**Contribution to Project:**

As part of her Ph.D. thesis, Lulu worked on another important topic of this project. She worked on developing efficient experimental designs for developing robust products and processes. The main idea is to propose a Bayesian optimal design criterion to generate single arrays that are much smaller in size than the most popular cross arrays. She is also working on developing and applying a Bayesian optimal design criterion to find optimal blocking in fractional factorial experiments. She was an intern at Proctor & Gamble during summer 2007 and motivated by a real problem in that industry, she is also working on developing design and analysis methods for mixture-of-mixture experiments. Lulu has presented her work in various conferences. She is fully supported by the grant.

Name: Ba, Shan**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Mr. Shan Ba is a Ph.D. student in ISYE. He did research on multi-layer designs and coupled Gaussian process models for computer experiments.

Name: Yan, Huan**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Mr. Huan Yan is a first year Ph.D. student in ISYE. He is learning the basics of the project and is expected to contribute more in

the coming years.

Undergraduate Student

Name: Wolf, Kyle

Worked for more than 160 Hours: Yes

Contribution to Project:

Kyle wolf worked on developing some class room experiments. He identified some good experiments, conducted them, and analyzed the results. These experiments are now implemented in teaching an undergraduate course on quality improvement.

Name: Schultz, Laura

Worked for more than 160 Hours: Yes

Contribution to Project:

Laura worked on analyzing several real case studies on robust parameter design.

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Other Collaborators or Contacts

T. Dasugpta, Y. Hung, C. F. J. Wu, Z. L. Wang, M. Yuan, S. L. Jeng, I. T. Yu, M. Ai.

Activities and Findings

Research and Education Activities:

Research activities:

Seventeen (17) papers are published, one (1) paper is accepted for publication, and two (2) papers are under review in top journals in statistics and engineering. One of the papers won a Journal's best paper award and another paper won best student paper award. The following are the major research activities.

1. Developed a prior distribution for the linear model parameters in a two-level experiment using a Gaussian process functional prior.
2. Extended the approach to three and higher level experiments and developed prior distributions for dealing with quantitative and qualitative factors.
3. Developed new meta-modeling methods for computer experiments that overcome the problems with the existing methods.
4. Developed efficient experimental designs for computer experiments.
5. Developed new analysis methods for robust parameter design.
6. Applied experimental design techniques for the synthesis of nanostructures.
7. Developed efficient methods for the analysis of experiments.
8. Developed Bayesian single arrays for studying robustness (which won the best student paper award from QSR section of INFORMS).
9. Developed a method for finding optimal specifications of degrading product characteristics (which won the best paper award from IIE transactions).
10. Developed a new modeling approach for mixture-of-mixture experiments and applied it to develop a new Pringles potato crisp at Procter and Gamble.

Education activities:

1. The topic of robust parameter design is introduced to the undergraduate students through the course ISyE 3039: Methods for quality

improvement. The computer experiments and Bayesian methods for the analysis of experiments are taught to the graduate students through the courses ISyE 7400: Advanced design of experiments and ISyE 6420 Bayesian Statistics.

2. Four Ph.D. students are partially supported by the grant; one graduated in Summer 2006 and another in Summer 2010.
3. Two undergraduate students worked on this project. One of them developed some classroom experiments, which are now implemented in the course ISyE 3039: Methods for quality improvement.
4. The findings are presented in various national and international conferences.

Findings:

The findings are described in the order of publications reported in the Journal Publications section.

1. A very simple and elegant approach is developed for obtaining the prior distribution for 2-level experiments. Its advantages are demonstrated through the analysis of several experiments. A new class of design criteria is proposed and their connections with the popular minimum aberration criterion are established. The work is published in *Technometrics*.
2. A new kriging method known as limit kriging is developed. This new predictor gives a better performance over the existing predictor when the constant mean assumption in the kriging model is unreasonable. Moreover, the new predictor seems to be robust to the misspecifications in the correlation parameters. The work is published in *Technometrics*. It was the lead article of the issue.
3. The approach in paper#1 is extended to the three and higher level experiments. The extension was not trivial because of many issues involved in higher level experiments that are not present in two-level experiments. For example, the type of factor, the type of correlation function, the type of coding scheme, the mixed-level nature of the experiments, etc. become important when dealing with higher level experiments, but are irrelevant for two-level experiments. The superiority of the proposed approach over the existing approaches is demonstrated using the analysis of some real experiments. This work forms a major part of James Delaney's Ph.D. thesis. The work is published in *Technometrics* and was the lead article of the issue.
4. Latin hypercube designs are quite popular in computer experiments. This work develops an approach to find optimal Latin hypercube designs by maximizing the inter-site distances between the points and minimizing the pair-wise correlation between the factors. The work is published in *Statistica Sinica*.
5. A new criterion for selecting adjustment factors in robust parameter design is proposed, which resolves many of the controversies surrounding Taguchi's approach. The work is published in *IIE Transactions on quality and reliability engineering*.
6. Experimental design techniques were applied for the synthesis of CdSe nanostructures. Robust parameter design techniques were applied to find processing conditions that produces nanostructures with consistent and high yield. The work is published in the *Journal of American Statistical Association*.
7. This work develops a fast algorithm for variable selection in designed experiments using a modification of the least angle regression technique. This new algorithm incorporates the effect hierarchy and heredity principles in variable selection and thus is capable of finding models that are more interpretable. The work is published in *Technometrics*.
8. Failure amplification method is used for improving the process capability when the failure probability is small. This work develops new methods for the modeling and analysis for failure amplification method. The work is published in *Journal of Quality Technology*.
9. This work develops a method for finding optimal specifications for quality characteristics that degrade over time. The main idea is to shift the manufacturing target of the characteristic so that the lifetime is increased with some sacrifice on quality. The work is published in *IIE Transactions on quality and reliability engineering*. It was a highlighted research article in the *Industrial Engineer Magazine* 40(2), page 57, 2008. The paper won the Best Paper Award from *IIE Transactions on quality and reliability engineering*.
10. The typical practice for analyzing industrial experiments is to identify statistically significant effects with a 5% level of significance and then to optimize the model containing only those effects. This work illustrates the danger in utilizing this approach and proposes alternative methodology using the practical significance level and empirical Bayes methods. The work is published in the *Journal of Quality Technology*.
11. Kriging is a useful method for developing metamodels for product design optimization. The most popular kriging method, known as ordinary kriging, uses a constant mean in the model. In this article, a modified kriging method is proposed, which has an unknown mean model.

Therefore it is called blind kriging. The unknown mean model is identified from experimental data using a Bayesian variable selection technique developed in paper # 1 and 3. Many examples from Ford motor company are presented which show remarkable improvement in prediction using blind kriging over ordinary kriging. This work is partially supported by the grant and is published in the ASME Journal of Mechanical Design.

12. It is critical to estimate control-by-noise interactions in robust parameter design. This can be achieved by using a cross array, which is a cross product of a design for control factors and another design for noise factors. However, the total run size of such arrays can be prohibitively large. To reduce the run size, single arrays are proposed in the literature, where a modified effect hierarchy principle is used for the optimal selection of the arrays. In this work, we argue that effect hierarchy is a property of the system and cannot be altered depending on the objective of an experiment. We propose a Bayesian approach to develop single arrays which incorporate the importance of control-by-noise interactions without altering the effect hierarchy. The approach is very general and places no restrictions on the number of runs or levels or type of factors or type of designs. A modified exchange algorithm is proposed for finding the optimal single arrays. We also explain how to design experiments with internal noise factors. This work has appeared in *Technometrics*. My student, Ms. Lulu Kang, won the best student paper award for this work from the Quality, Statistics, and Reliability section of INFORMS at the annual meeting in San Diego, 2009.

13. In this work, we consider structured variable selection and estimation, which accounts for hierarchical structural relationship among predictors in linear regression. We propose nonnegative garrote methods that can naturally incorporate such relationships defined through effect heredity principles in variable selection and estimation. We show that the methods are very easy to compute and enjoy nice theoretical properties. We also show that the methods can be easily extended to deal with more general regression problems such as generalized linear models. The work is published in the *Annals of Applied Statistics*.

14. It is not easy to provide an ordering for the effects involving two-level factors and the components of four-level factors. The existing ordering is adhoc and therefore, the validity of minimum aberration designs for mixed two- and four-level factors can be challenged. In this work, we propose a Bayesian approach to overcome this problem. We use functionally induced priors developed in Papers# 1 and 3 so that the effects can be ordered less ambiguously. Using the Bayesian approach we develop an optimal design criterion that leads to a new minimum aberration type criterion. Different from the existing work, this approach is also capable of distinguishing between qualitative and quantitative factors. Many examples are given to demonstrate the advantages of the proposed approach. A catalogue of optimal designs is also provided. The work is published in *Biometrika*.

15. The presence of block effects makes the optimal selection of fractional factorial designs a difficult task. The existing frequentist methods try to combine treatment and block wordlength patterns and apply minimum aberration criterion to find the optimal design. However, ambiguities exist in combining the two wordlength patterns and therefore, the optimality of such designs can be challenged. A Bayesian approach is proposed to overcome this problem. The main technique is to postulate a model and a prior distribution to satisfy the common assumptions in blocking and then, to develop an optimal design criterion for the efficient estimation of treatment effects. We apply our method to develop regular, nonregular, and mixed-level blocked designs. Several examples are presented to illustrate the advantages of the proposed method. The paper is published in the *Journal of Statistical Planning and Inference*.

16. In mixture-of-mixture experiments, major components are defined as the components which themselves are mixtures of some other components, called minor components. Sometimes, components are divided into different categories, where each category is called a major component, and the components within a major component become minor components. The special structure of the mixture-of-mixture experiment makes the design and modeling approaches different from a typical mixture experiment. In this paper, we propose a new model called major-minor model, which overcomes some of the limitations of the commonly used multiple-Scheffe model. We also provide a strategy for designing the experiments. We then apply the proposed design and modeling approach to a mixture-of-mixture experiment conducted to formulate a new Pringles potato crisp. The paper is published in *Technometrics*.

17. This work is motivated by a problem of optimizing printed circuit board manufacturing using design of experiments. The data is binary, which poses challenges in model fitting and optimization. We use the idea of failure amplification method to increase the information supplied by the data and then use a Bayesian approach for model fitting. The Bayesian approach is implemented using Gaussian process models on a latent variable representation. It is demonstrated that the Bayesian approach is highly suitable for optimizing a process with binary data. The paper is published in *Applied Stochastic Models for Business and Industry*.

18. Space-filling designs such as Latin hypercube designs (LHDs) are widely used in computer experiments. However, finding an optimal LHD with good space-filling properties is computationally cumbersome. On the other hand, the well-established factorial designs in physical experiments are unsuitable for computer experiments owing to the redundancy of design points when projected onto a subset of factor space. In this work, we developed a new class of space-filling designs developed by splitting two-level factorial designs into multiple layers. The method takes advantages of many available results in factorial design theory and therefore, the proposed Multi-layer designs (MLDs) are easy to generate. Moreover, our numerical study shows that MLDs can have better space-filling properties than optimal LHDs. The paper is accepted

for publication in the Journal of the American Statistical Association.

19. A more accurate approach for approximating computationally expensive functions is proposed by coupling two Gaussian process models together. Different from the existing kriging method, the new predictor incorporates both the non-stationarity in mean and variance and is more capable of approximating surfaces that are not second-order stationary. It also gives better prediction performance when the experimental design is unable to cover the region well, and can improve the prediction intervals by quantifying the change of variability associated with the response. The paper is under review in the Annals of Applied Statistics.

20. In this work we propose a new robust regression method called regression with outlier shrinkage (ROS). It improves over the other robust regression methods such as Least Trimmed Squares (LTS) in the sense that it can achieve maximum breakdown value and full asymptotic efficiency simultaneously. Moreover, its computational complexity is no more than that of LTS. We also propose a sparse estimator, called sparse regression with outlier shrinkage (SROS), for robust variable selection and estimation. It is proven that SROS can not only give consistent selection but also estimate the nonzero coefficients with full asymptotic efficiency under the normal model. In addition, we introduce a concept of nearly regression equivariant estimator for understanding the breakdown properties of sparse estimators, and prove that SROS achieves the maximum breakdown value of nearly regression equivariant estimators. Numerical examples are presented to illustrate our methods. The paper is under review in the Journal of the American Statistical Association.

Training and Development:

Four graduate students worked on the project. One of the important topics of the project formed the major part of Mr. Delaney's Ph.D. thesis. He developed fundamental understanding of Bayesian methods and its applications to the design and analysis of experiments. His work will be a major contribution to the field. He graduated in Summer 2006 and presently working as an assistant professor in Temple University. Ms. Kang worked on developing Bayesian single arrays for robust parameter design, Bayesian optimal blocking, the modeling and analysis of mixture-of-mixture experiments, and computer experiments. She graduated in Summer 2010 and is now working as an assistant professor in Illinois Institute of Technology. Mr. Ba worked on new design and analysis methods for computer experiments. Mr. Yan is a first year Ph.D. student and is expected to work on projects related to nanotechnology. Two undergraduate students, Mr. Wolf and Ms. Schultz, also worked on this project.

Outreach Activities:

The finding of the project are written as papers and submitted to reputed journals for publication. The work is presented at various conferences.

The PI presented the findings of the project at the International conference on design of experiments, Memphis, 2005; Design and analysis of experiments conference, Santa Fe, 2005; INFORMS Annual meeting, San Francisco, 2005 (two presentations); Spring research conference, Park City, 2005; Joint research conference, Knoxville, 2006; INFORMS Annual meeting, Pittsburgh, 2006; Department of Statistics, University of Georgia, Athens, 2006, SQC&OR unit, Indian Statistical Institute, Calcutta, 2007; INFORMS Annual meeting, Seattle, 2007, CMMI grantee conference, Knoxville, 2008, Quality and Productivity Research Conference, Madison, 2008; CMMI grantee conference, Hawaii, 2009; INFORMS Annual meeting, San Diego, 2009; IMS Annual meeting, Gothenberg, Sweden, 2010; CMMI grantee conference, Atlanta, 2011; and Isaac Newton Institute workshop on design of experiments, Cambridge, UK, 2011.

James Delaney presented the work at INFORMS Annual conference, San Francisco, 2005; Joint research conference, Knoxville, 2006; INFORMS Annual conference, Pittsburgh, 2006; Design of experiments conference, Southampton, 2006; and DMI grantee conference, St. Louis, 2006.

Lulu Kang presented the work at the INFORMS annual meeting, Seattle, 2007; CMMI grantee conference, Knoxville, 2008; Spring Research Conference, Atlanta, 2008; Quality and Productivity Research Conference, Madison, 2008; Joint Statistical Meeting, Denver, 2008; and INFORMS Annual meeting, Washington D.C., 2008; Spring Research Conference, Vancouver, Canada, 2009; CMMI grantee conference, Hawaii, 2009; and Joint Statistical Meeting, Washington, D. C., 2009. She presented the work in INFORMS Annual meeting, San Diego, 2009 and won the best student paper competition in the QSR section.

The PI organized seven sessions in various national/international conferences for increasing the public awareness of the project activities.

Journal Publications

V. Roshan Joseph, "A Bayesian Approach to the Design and Analysis of Fractionated Experiments", Technometrics, p. 219, vol. 48, (2006). Published,

- V. Roshan Joseph, "Limit Kriging", *Technometrics*, p. 458, vol. 48, (2006). Published,
- V. Roshan Joseph and James D. Delaney, "Functionally Induced Priors for the Analysis of Experiments", *Technometrics*, p. 1, vol. 49, (2007). Published,
- V. Roshan Joseph and Hung, Y., "Orthogonal-Maximin Latin Hypercube Designs", *Statistica Sinica*, p. 171, vol. 18, (2008). Published,
- V. Roshan Joseph, "Taguchi's Approach to Robust Parameter Design: A New Perspective", *IIE Transactions*, p. 805, vol. 39, (2007). Published,
- Dasgupta, T., Ma, C., V. Roshan Joseph, Wang, Z. L., and Wu, C. F. J., "Statistical Modeling and Analysis for Robust Synthesis of Nanostructures", *Journal of the American Statistical Association*, p. 594, vol. 103, (2008). Published,
- Yuan, M., V. Roshan Joseph, and Lin, Y., "An Efficient Variable Selection Approach for Analyzing Designed Experiments", *Technometrics*, p. 430, vol. 49, (2007). Published,
- Jeng, S-L., V. Roshan Joseph, and Wu, C. F. J., "Modeling and Analysis Strategies for Failure Amplification Method", *Journal of Quality Technology*, p. 128, vol. 40, (2008). Published,
- V. Roshan Joseph and Yu, I.-T., "Optimal Specifications for Degrading Characteristics", *IIE Transactions*, p. 102, vol. 40, (2008). Published,
- V. Roshan Joseph and Delaney, J. D., "Analysis of Optimization Experiments", *Journal of Quality Technology*, p. 282, vol. 40, (2008). Published,
- V. Roshan Joseph, Hung, Y., and Sudjianto, A., "Blind Kriging: A New Method for Developing Metamodels", *ASME Journal of Mechanical Design*, p. 031102-1, vol. 130, (2008). Published,
- Kang, L. and V. Roshan Joseph, "Bayesian Optimal Single Arrays for Robust Parameter Design", *Technometrics*, p. , vol. 51, (2009). Published,
- Yuan, M., V. Roshan Joseph, and Zou, H., "Structured Variable Selection and Estimation", *Annals of Applied Statistics*, p. 1738, vol. 3, (2009). Published,
- V. Roshan Joseph, Ai, M., and Wu, C. F. J., "Bayesian-Inspired Mixed Two- and Four-Level Designs", *Biometrika*, p. 95, vol. 96, (2009). Published,
- Ai, M., Kang, L., and V. Roshan Joseph, "Bayesian Optimal Blocking of Factorial Designs", *Journal of Statistical Planning and Inference*, p. 3319, vol. 139, (2009). Published,
- Kang, L., V. Roshan Joseph, and Brenneman, W., "Design and Modeling Strategies for Mixture-of-Mixture Experiments", *Technometrics*, p. 125, vol. 53, (2011). Published,
- Yu, I.-T and V. Roshan Joseph, "Bayesian Process Optimization Using Binary Data", *Applied Stochastic Models for Business and Industry*, p. 402, vol. 27, (2011). Published,
- Ba, S. and V. Roshan Joseph, "Multi-Layer Designs for Computer Experiments", *Journal of the American Statistical Association*, p. , vol. , (2011). Accepted,
- Ba, S. and V. Roshan Joseph, "Coupled Gaussian Process Models", *Annals of Applied Statistics*, p. , vol. , (2011). Submitted,
- Xiong, S. and V. Roshan Joseph, "Regression with Outlier Shrinkage", *Journal of the American Statistical Association*, p. , vol. , (2011). Submitted,

Books or Other One-time Publications**Web/Internet Site****Other Specific Products****Contributions****Contributions within Discipline:**

The most common practice in the field of quality engineering is to use only the experimental data to make conclusions. The Bayesian methods allow one to use prior knowledge and thus, the information from the experimental data is combined with the prior knowledge to make conclusions. The existing Bayesian methods for experiments are difficult to use. A major contribution of this project is in the development of an intuitive and easy-to-use Bayesian approach to experiments.

One of the main difficulties with the Bayesian approach is to specify a prior distribution for the large number of parameters in the statistical model. This project developed a methodology to induce the prior distribution from a prior on the underlying transfer function, which can be easily done. The induced prior satisfies the well-known principles of effect hierarchy and effect heredity and thus will appeal to both frequentists and Bayesians. The Bayesian optimal design criterion has some connections with the popular frequentist criterion known as minimum aberration, which is established for the first time in the literature.

Another major contribution of the project is to the field of computer experiments. The project developed a new experimental design known as orthogonal-maximin latin hypercube for the design of computer experiments and a new metamodeling technique known as limit kriging for the analysis of computer experiments. A new class of designs called multi-layer designs is also developed.

The project also developed new methods for the design and analysis of experiments for developing robust products and processes. The proposed Bayesian single array is much smaller than the usual experimental designs used for studying robustness and thus, has the potential to save time and money for the investigator.

Contributions to Other Disciplines:

The findings from this project are likely to contribute to all disciplines of science and engineering that use experimental investigation. We have carried out successful experiments for improving the yield of certain nano materials and for developing products in industries.

Contributions to Human Resource Development:**Contributions to Resources for Research and Education:****Contributions Beyond Science and Engineering:****Conference Proceedings****Categories for which nothing is reported:**

Organizational Partners

Any Book

Any Web/Internet Site

Any Product

Contributions: To Any Human Resource Development

Contributions: To Any Resources for Research and Education

Contributions: To Any Beyond Science and Engineering

Any Conference



Georgia Institute of Technology

The H. Milton Stewart School
of Industrial & Systems Engineering

June 28, 2011

Dr. Russell R. Barton
Program Director
Manufacturing Enterprise Systems
Division of Civil, Mechanical, and Manufacturing Innovation
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230

Dear Dr. Barton:

I am very pleased to see the progress of Dr. Roshan Joseph Vengazhiyil's project "CAREER: Design and Analysis of Experiments for Developing Robust Products and Processes". His research and education plans are of significant interest to our School.

I reaffirm that the PI's career development plan is supported by, and integrated into, the educational and research goals of the School and the Institute. I am personally committed to the support and professional development of the PI.

Sincerely,

Michael E. Thomas
Interim Chair and Professor